University of Toronto
Department of Economics

Working Paper 532

(Partially) Identifying potential outcome distributions in triangular systems

By Ismael Mourifie and Yuanyuan Wan

January 29, 2015

(PARTIALLY) IDENTIFYING POTENTIAL OUTCOME DISTRIBUTIONS IN TRIANGULAR SYSTEMS

ISMAEL MOURIFIÉ† AND YUANYUAN WAN‡

DEPARTMENT OF ECONOMICS   UNIVERSITY OF TORONTO

ABSTRACT. In this paper we propose a new unifying approach to (partially) identify potential outcome distributions in a non-separable triangular model with a binary endogenous variable and a binary instrument. Our identification strategy provides a testable condition under which the objects of interest are point identified. When point identification is not achieved, we provide sharp bounds on the potential outcome distributions and the difference of marginal distributions.

Keywords: Potential outcomes, triangular system, point and partial identification, sharp bounds.

This version: Monday 29th December, 2014

1. INTRODUCTION

This paper studies a non-separable triangular model with a binary endogenous variable and a binary instrument

$$\begin{cases} Y = g(D, U), \\ D = \mathbf{1}\{V \leq p(Z)\}, \end{cases} \tag{1}$$

where $Y \in \mathcal{Y} \subseteq \mathbb{R}$ is the outcome variable, $D \in \{0,1\}$ is the endogenous regressor (treatment), and $Z \in \{0,1\}$ is the binary instrument. (Y, D, Z) are observables. g and p are unknown functions with $g(d, \cdot)$ nondecreasing and left-continuous for $d \in \{0,1\}$. U and V are scalar-valued latent variables. Additionally U and V are normalized to be uniformly distributed i.e., $U, V \sim \text{Uniform}[0,1]$. Assuming that Z is independent with (U, V),[1] as will be discussed in more detail, we derive sharp bounds on the potential outcome distributions

$$\mathbb{P}(Y_d \leq y), d \in \{0,1\}$$

† Corresponding author. Department of Economics, University of Toronto, 150 St. George Street, Toronto ON M5S 3G7, Canada. ismael.mourifie@utoronto.ca.

‡yuanyuan.wan@utoronto.ca.

[1]In the presence of the exogenous covariate X, this assumption is strengthened. X is left out to simplify the notation, and its addition will be discussed in Section 3.

and the difference in marginal distributions

$$\mathbb{P}(Y_1 \leq y) - \mathbb{P}(Y_0 \leq y)$$

of the whole population. Identification of these quantities are especially important for the analyses of heterogeneous treatment effects. As mentioned by Imbens and Rubin (1997) and Angrist and Pischke (2008), these distributions are useful for policy makers who want to take into account differences in the dispersion of earnings when contemplating the merits of one program or treatment versus another.

Our main contribution is to provide a new unifying approach to (partially) identify potential outcome distributions in this setup. When g is strictly monotone in U (Y is therefore continuously distributed), bounds collapse to a point and the point identification result achieved by Vuong and Xu (2014) is recovered. When the outcome variable Y is binary, bounds are exactly as those found by Shaikh and Vytlacil (2011). This identification strategy allows a testable condition to be derived under which objects of interest are point identified. This testable condition reveals that identification can be achieved, even if g is weakly monotone in U, which encompasses the cases where Y is either censored, truncated, discrete, or mixed continuous-discrete outcomes.

A model similar to Equation (1) has also been studied by Vytlacil and Yildiz (2007). The identification strategy in their study requires the existence of an additional exogenous covariate and provides rank conditions based on exogenous covariates, under which it is possible to identify the average effect. Recently, Vuong and Xu (2014) generalized Vytlacil and Yildiz (2007)'s rank condition to point identify the quantile functions. Both papers do not discuss partial identification when the proposed rank condition fails to hold. Our paper complements this research by providing sharp bounds on potential outcome distributions whenever rank condition fails to hold. When the rank condition holds our bounds coincide with the identification results of these studies.

The rest of the paper is organized as follows. Section 2 presents the identification strategy. Section 3 generalizes the method to the case where additional exogenous covariates are available. Proofs are collected in Appendix.

2. Identification strategy

We make the following assumptions:

Assumption 1. $(U, V) \perp Z$.

Assumption 2. *$g(d,\cdot)$ is non-decreasing and left-continuous for $d \in \{0,1\}$.*

Here $p(z) = \mathbb{P}(D = 1|Z = z)$, and without a loss of generality, we assume that $p(1) \geq p(0)$.[2] Assumption 1 is a common assumption in the literature. Assumption 2 requires weak monotonicity of $g(d,\cdot)$ as a function of U, for each $d \in \{0,1\}$. There are no restrictions on the relative ranking of $g(1,\cdot)$ and $g(0,\cdot)$ i.e., both can cross each other many times.

Before presenting formal results, a heuristic argument for identification strategy will first be provided. Under Assumption 1, the distribution function of Y_1 can be decomposed as follows:

$$\mathbb{P}(Y_1 \leq y) \quad = \quad \mathbb{P}(Y_1 \leq y, V \leq p(1)|Z = 1) + \mathbb{P}(Y_1 \leq y, V > p(1)|Z = 1),$$

where the first right hand side term $\mathbb{P}(Y_1 \leq y, V \leq p(1)|Z = 1) = \mathbb{P}(Y \leq y, D = 1|Z = 1)$ is identified from the data. The second right hand side term $\mathbb{P}(Y_1 \leq y, V > p(1)|Z = 1) = \mathbb{P}(Y_1 \leq y, D = 0|Z = 1)$ is the unobserved counterfactual.

In this paper, the sharp bounds for $\mathbb{P}(Y_1 \leq y, V > p(1)|Z = 1)$ are derived by taking advantage of the "identified distribution of potential outcomes for compliers" i.e., $\mathbb{P}(Y_d \leq y|c)$ for $d = 0, 1$. In fact, $\mathbb{P}(Y_1 \leq y, V > p(1)|Z = 1)$ is a weighting function of the distribution of Y_1 for the "defiers" and "never-takers" in the language of Imbens and Angrist (1994). Since model (1) imposes a monotonicity restriction on the treatment, it rules out the existence of the "defiers" (See Vytlacil, 2002). Therefore, the unobserved quantity involves only the "never-takers". As suggested by the name, "never-takers" never take the treatment, and the potential outcome under the treatment, i.e., Y_1, cannot be observed from the data. However, Imbens and Rubin (1997) show that the potential outcome distribution of the compliers for treated $\mathbb{P}(Y_1 \leq y|c)$ and untreated $\mathbb{P}(Y_0 \leq y|c)$ are identifiable from the data. These distributions of the compliers are used as the matching function criteria in this study. Indeed, for a fixed $y \in \mathcal{Y}$, if there exists a y' that makes the compliers indifferent to being in the treated versus the untreated group, in the sense that

$$\Delta(y', y) \equiv \mathbb{P}(Y_0 \leq y'|c) - \mathbb{P}(Y_1 \leq y|c) = 0,$$

[2] The model described in Equation (1) is related to potential outcome models. In particular, one can define D_j, $j = 1, 2$ as a potential treatment when the value of Z is externally set to j. Likewise, $Y_j = g(j, U)$ can be defined as the potential outcome when D is set to j externally (in potential outcome models, U does not need to be scalar-valued). In this notation, $Y = Y_1 D + Y_0(1 - D)$ and $D = D_1 Z + D_0(1 - Z)$.

then this value y' makes also "never-takers" indifferent to being treated versus untreated. Because the potential outcome distribution of Y_0 for never-takers is identified, then the identification $\mathbb{P}(Y_1 \leq y, D = 0|Z = 1) = \mathbb{P}(Y_0 \leq y', D = 0|Z = 1)$ is achieved. If a "perfect" match does not exist $\Delta(y', y) > 0$ or $\Delta(y', y) < 0$ then $\mathbb{P}(Y_1 \leq y, D = 0|Z = 1)$ can be bound from below or above, depending on the sign of $\Delta(y', y)$.

The quantity $\Delta(y', y)$ plays an important role in the identification strategy, and $\Delta(y', y)$ is identified (See Imbens and Rubin, 1997). In particular,

$$\Delta(y', y) = \frac{H(y', y)}{p(1) - p(0)}$$

where

$$\begin{aligned} H(y', y) \equiv & \Big[\mathbb{P}(Y \leq y', D = 0|Z = 0) - \mathbb{P}(Y \leq y', D = 0|Z = 1)\Big] \\ & - \Big[\mathbb{P}(Y \leq y, D = 1|Z = 1) - \mathbb{P}(Y \leq y, D = 1|Z = 0)\Big] \\ & = \mathbb{P}(Y_0 \leq y', p(0) < V \leq p(1)) - \mathbb{P}(Y_1 \leq y, p(0) < V \leq p(1)). \end{aligned} \tag{2}$$

We will refer to H as " the matching function" hereafter. The above discussion is summarized in the following lemma.

Lemma 1. *Let g_d^{-1} be the generalized inverse of g_d. Under Assumptions 1 and 2 the following occur:*

(1) $sign(H(y', y)) = sign([g_0^{-1}(y') - g_1^{-1}(y)])$ *where* $sign(a) = 1\{a > 0\} - 1\{a < 0\}$.

(2) *For a fixed y (or y'), $H(y', y)$ is a non-decreasing (or non-increasing) function in y' (or y').*

The proof of the lemma follows straightforwardly from Equation (2) and is therefore omitted. From Lemma 1-(1), and because $\mathbb{P}(Y_1 \leq y, D = 0|Z = z) = \mathbb{P}(U \leq g_1^{-1}(y), D = 0|Z = z)$, there is the following. For each given y:

$$\mathbb{P}(U \leq g_1^{-1}(y), D = 0|Z = z) \begin{cases} \leq \mathbb{P}(U \leq g_0^{-1}(y'), D = 0|Z = z) \text{ if } sign(H(y', y)) = 1 \\ = \mathbb{P}(U \leq g_0^{-1}(y'), D = 0|Z = z) \text{ if } sign(H(y', y)) = 0 \\ \geq \mathbb{P}(U \leq g_0^{-1}(y'), D = 0|Z = z) \text{ if } sign(H(y', y)) = -1, \end{cases}$$

where $\mathbb{P}(U \leq g_0^{-1}(y'), D = 0|Z = z) = \mathbb{P}(Y \leq y', D = 0|Z = z)$ is the observed factual of the untreated.

To bound the unobserved counterfactual, previous research has used variations in the instrument, across treatment, or from the exogenous covariate, when it is available. Here we show that variation in the dependant outcome Y can also be used. Before stating the primary result, the following sets are defined: The focus is put on the upper bound to deliver the main idea, and similar results hold for

TABLE 1. Collection of Sets

$\Omega_{01}^{+}(y) \equiv \{y' : H(y', y) \geq 0\}$	$\Omega_{01}^{-}(y) \equiv \{y' : H(y', y) \leq 0\}$
$\Delta_{01}^{+}(y) \equiv \{y' : H(y, y') \geq 0\}$	$\Delta_{01}^{-}(y) \equiv \{y' : H(y, y') \leq 0\}$

the lower bound. Notice that by construction, we have

$$\mathbb{P}(Y_1 \leq y, D = 0|Z = z) \leq \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0|Z = z).$$

Thus,

$$\mathbb{P}(Y_1 \leq y) \leq \inf_{z \in \{0,1\}} \left[\mathbb{P}(Y \leq y, D = 1|Z = z) + \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0|Z = z) \right].$$

As shown in Claim 3 in the appendix, the upper bound can be simplified as:

$$\begin{aligned} &\mathbb{P}(Y \leq y, D = 1|Z = 1) + \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0|Z = 1) \\ &\qquad \leq \mathbb{P}(Y \leq y, D = 1|Z = 0) + \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0|Z = 0). \end{aligned} \tag{3}$$

This simplification is helpful for the inference, and similar simplification has already been discussed in the literature. Heckman and Vytlacil (2001) explained that the monotonicity imposed on the treatment does not improve the Manski (1990) bounds, but it does provide a simplification of the bounds. In addition to the monotonicity of the treatment, Assumption 2 allows for further simplification of the bounds. The main identification result is given in the following theorem.

Theorem 1. *Let* (Y,D,Z)*,* $g_d(.)$*, and* (U,V) *define the triangular system (1). Under assumptions (1) and (2) the following bounds are sharp:*

$$\mathbb{P}(Y\le y,D=1|Z=1)+\sup_{y'\in\Omega_{01}^{-}(y)}\mathbb{P}(Y\le y',D=0|Z=1)\le\mathbb{P}(Y_1\le y)\le$$
$$\mathbb{P}(Y\le y,D=1|Z=1)+\inf_{y'\in\Omega_{01}^{+}(y)}\mathbb{P}(Y\le y',D=0|Z=1),$$

and

$$\mathbb{P}(Y\le y,D=0|Z=0)+\sup_{y'\in\Delta_{01}^{+}(y)}\mathbb{P}(Y\le y',D=1|Z=0)\le\mathbb{P}(Y_0\le y)\le$$
$$\mathbb{P}(Y\le y,D=0|Z=0)+\inf_{y'\in\Delta_{01}^{-}(y)}\mathbb{P}(Y\le y',D=1|Z=0).$$

Proof. See Appendix A.

Remark 1. *When* Y *is binary and without additional exogenous covariates, the bounds are exactly as the bounds derived by Shaikh and Vytlacil (2011). When* g_d *is strictly increased in* U*, the bounds collapse to a point and then the identification result of Vuong and Xu (2014) is recovered. Indeed, when* g_d *is strictly increased,* y' *can be found such that* $H(y',y)=0$*, therefore identification is always obtained in this particular case.*

Remark 2. *Our identification strategy cannot be directly applied in the model with sector-specific heterogeneity i.e.,* $Y_d=g_d(U_d)$*, where* $U_1\neq U_0$*. However, this methodology holds if we impose the rank similarity assumption discussed in Chernozhukov and Hansen (2005) i.e.,* $U_1|V\sim U_0|V$*.*

Remark 3. *In the proof, a joint distribution is constructed based on observed distributions that respects all restrictions imposed on the model that allows* $\mathbb{P}(Y_0\le y)$ *to equals his lower bound and* $\mathbb{P}(Y_1\le y)$ *his upper bound. Therefore, the bounds on the difference in marginal distributions obtained by simple subtraction of the bounds on the potential outcome distribution are also sharp.*

The presence of a multi-valued instrument strengthens the identification power of this strategy. A general matching function can be defined as

$$\begin{aligned} H(y',y,z',z) \quad \equiv \quad & \Big[\mathbb{P}(Y \le y', D=0|Z=z) - \mathbb{P}(Y \le y', D=0|Z=z')\Big] \\ & -\Big[\mathbb{P}(Y \le y, D=1|Z=z') - \mathbb{P}(Y \le y, D=1|Z=z)\Big]. \end{aligned}$$

If y' that allows $H(y',y,z',z)=0$ cannot be found, then a y' that allows $H(y',y,z'',z')=0$ should be found.

3. Generalization

Let us consider an extension with additional exogenous covariates X that enter both equations in model (1) in the sense that

$$\begin{cases} Y = g(D,X,U), \\ D = \mathbf{1}\{V \le p(X,Z)\}, \end{cases} \tag{4}$$

with $(U,V) \perp (Z,X)$. For sake of simplicity, let us assume that

Assumption 3. *$Supp(p(X,Z),X) = Supp(p(X,Z)) \times Supp(X)$, where Supp denotes the support.*

This assumption is not required for a partial identification approach but is used here to simplify the notation. Vytlacil and Yildiz (2007) and Vuong and Xu (2014) require that $Supp(p(X,Z)|X = x) \cap Supp(p(X,Z)|X = x')$ contains at least two values. Whenever Z and X are binary, it is equivalent to Assumption 3. Shaikh and Vytlacil (2011) also use Assumption 3 to provide sharp bounds on the average treatment effect, when Y is binary. Mourifié (2013) explains how sharp bounds can be obtained when Assumption 3 fails to hold. Therefore, if Assumption 3 does not hold Mourifié (2013)'s approach can be easily extended to the present context. A generalized matching function is defined as:

$$\begin{aligned} H(y',y,x',x) \quad \equiv \quad & \Big[\mathbb{P}(Y \le y', D=0|X=x', Z=0) - \mathbb{P}(Y \le y', D=0|X=x', Z=1)\Big] \\ & -\Big[\mathbb{P}(Y \le y, D=1|X=x, Z=1) - \mathbb{P}(Y \le y, D=1|X=x, Z=0)\Big]. \end{aligned}$$

By adapting the derivation from the latter section, we can derive the following sharp bounds:

$$\mathbb{P}(Y \leq y, D = 1 | X = x, Z = 1) + \sup_{(y',x') \in \Omega_{01}^{-}(y,x)} \mathbb{P}(Y \leq y', D = 0 | X = x', Z = 1)$$
$$\leq \mathbb{P}(Y_1 \leq y | X = x) \leq$$
$$\mathbb{P}(Y \leq y, D = 1 | X = x, Z = 1) + \inf_{(y',x') \in \Omega_{01}^{+}(y,x)} \mathbb{P}(Y \leq y', D = 0 | X = x', Z = 1),$$

where $\Omega_{01}^{+}(y,x) \equiv \{y' : H(y', y, x', x) \geq 0\}$ and $\Omega_{01}^{-}(y,x) \equiv \{y' : H(y', y, x', x) \leq 0\}$. Sharp bounds on $\mathbb{P}(Y_0 \leq y | X = x)$ can be similarly derived. Identification is achieved when x' exists such that $H(y, y, x', x) = 0$. This is equivalent to the rank condition proposed by Vytlacil and Yildiz (2007). If (y', x') exists in such a way to allow $H(y', y, x', x) = 0$, then the generalized rank condition proposed by Vuong and Xu (2014) is recovered. However, having such a perfect match cannot be achieved in many applications. In such a case the partial identification approach is useful because it provides sharp bounds on potential outcome distributions.

References

ANGRIST, J. D., AND J.-S. PISCHKE (2008): *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

CHERNOZHUKOV, V., AND C. HANSEN (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73(1), 245–261.

HECKMAN, J. J., AND E. J. VYTLACIL (2001): *Instrumental variables, selection models, and tight bounds on the average treatment effect*. Springer.

IMBENS, G. W., AND J. D. ANGRIST (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), 467–475.

IMBENS, G. W., AND D. B. RUBIN (1997): "Estimating Outcome Distributions for Compliers in Instrumental Variables Models," *The Review of Economic Studies*, 64(4), 555–574.

KITAGAWA, T. (2014): "A Test for Instrument Validity," Working paper.

MANSKI, C. F. (1990): "Nonparametric bounds on treatment effects," *The American Economic Review*, pp. 319–323.

MOURIFIÉ, I. (2013): "SHARP BOUNDS ON TREATMENT EFFECTS IN A BINARY TRIANGULAR SYSTEM," Discussion paper, University of Toronto, Department of Economics.

MOURIFIÉ, I., AND Y. WAN (2014): "Testing LATE Assumptions," *Working Paper*.

NELSEN, R. B. (2006): *An introduction to copulas*. Springer.

SHAIKH, A. M., AND E. J. VYTLACIL (2011): "Partial identification in triangular systems of equations with binary dependent variables," *Econometrica*, 79(3), 949–955.

VUONG, Q., AND H. XU (2014): "Counterfactual Mapping and Individual Treatment Effects in Nonseparable Models with Discrete Endogeneity," Discussion paper, Working Paper.

VYTLACIL, E. (2002): "Independence, Monotonicity, and Latent Index Models: An Equivalence Result," *Econometrica*, 70(1), 331–341.

VYTLACIL, E., AND N. YILDIZ (2007): "Dummy Endogenous Variables in Weakly Separable Models," *Econometrica*, 75(3), pp. 757–779.

APPENDIX A. PROOF OF THEOREM 1

Note first that the model implies the following restriction on observables (See Kitagawa, 2014; Mourifié and Wan, 2014, for details),

$$\mathbb{P}(Y \in A, D = 0|Z = 1) \leq \mathbb{P}(Y \in A, D = 0|Z = 0) \tag{5}$$

$$\mathbb{P}(Y \in A, D = 1|Z = 0) \leq \mathbb{P}(Y \in A, D = 1|Z = 1) \tag{6}$$

We propose the following joint distribution. For generic argument $(y,z) \in \mathcal{Y} \cup \{0,1\}$, Almost surely in Z,

$$\mathbb{P}(\tilde{g}_1(\tilde{U}) \leq y, \tilde{V} < p(z)|Z) = \mathbb{P}(Y \leq y, D = 1|Z = z), \tag{7}$$

$$\mathbb{P}(\tilde{g}_1(\tilde{U}) \leq y, \tilde{V} > p(z)|Z) = \mathbb{P}(Y \leq y, D = 1|Z = 1) \tag{8}$$

$$+ \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0|Z = 1) - \mathbb{P}(Y \leq y, D = 1|Z = z), \tag{9}$$

$$\mathbb{P}(\tilde{g}_0(\tilde{U}) \leq y, \tilde{V} > p(z)|Z) = \mathbb{P}(Y \leq y, D = 0|Z = z), \tag{10}$$

$$\mathbb{P}(\tilde{g}_0(\tilde{U}) \leq y, \tilde{V} < p(z)|Z) = \mathbb{P}(Y \leq y, D = 0|Z = 1) \tag{11}$$

$$+ \sup_{y' \in \Delta_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 1|Z = 1) - \mathbb{P}(Y \leq y, D = 0|Z = z), \tag{12}$$

$$\tilde{g}_1^{-1}(y) = \mathbb{P}(Y \leq y, D = 1|Z = 1) + \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0|Z = 1), \tag{13}$$

$$\tilde{g}_0^{-1}(y) = \mathbb{P}(Y \leq y, D = 0|Z = 1) + \sup_{y' \in \Delta_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 1|Z = 1), \tag{14}$$

where $p(z) = \mathbb{P}(D = 1|Z = z)$.

The theorem is proved by the following three steps.

A.1. **Step 1: Show that $\tilde{g}_d(y)$ are non-decreasing left-continuous.** We have to show here that $\tilde{g}_1^{-1}(y)$ is non-decreasing right-continuous function. Indeed, we can show that $\tilde{g}_1^{-1}(y)$ is a well defined CDF.

(1) $\tilde{g}_1^{-1}(y)$ is a non-decreasing function.
$y_1 < y_1' \Rightarrow \Omega_{01}^{+}(y_1') \subseteq \Omega_{01}^{+}(y_1) \Rightarrow \inf_{\tilde{y}_1 \in \Omega_{01}^{+}(y_1)} \mathbb{P}(Y \leq \tilde{y}_1, D = 0|Z = 1) \leq \inf_{\tilde{y}_1' \in \Omega_{01}^{+}(y_1')} \mathbb{P}(Y \leq \tilde{y}_1', D = 0|Z = 1) \Rightarrow \tilde{g}_1^{-1}(y_1) \leq \tilde{g}_1^{-1}(y_1')$.

(2) It is readily verifiable that $\lim_{y \to \underline{y}} \tilde{g}_1^{-1}(y) = 0$ and $\lim_{y \to \overline{y}} \tilde{g}_1^{-1}(y) = 1$ where $\underline{y}$ and $\overline{y}$ are the bounds of $\mathcal{Y}$.

(3) Right-continuous. Since the first term in the definition of $\tilde{g}_1^{-1}$ is right continuous by construction, it remains to verify that the second term is right continuous. Let y_n be a sequence such that $y_n \downarrow y$.

To simplify notation, define $A(y') = \mathbb{P}(Y \leq y', D = 0|Z = 1)$, $B(y') = \mathbb{P}(Y \leq y', D = 0|Z = 0) - \mathbb{P}(Y \leq y', D = 0|Z = 1)$, and $C(y) = \mathbb{P}(Y \leq y, D = 1|Z = 1) - \mathbb{P}(Y \leq y, D = 1|Z = 0)$. We want to show that $\inf_{y' \in \Omega_{01}^{+}(y_n)} A(y') \to \inf_{y' \in \Omega_{01}^{+}(y)} A(y')$, where

$$\Omega^+_{01}(y) = \{y' \in \mathcal{Y} : B(y') \geq C(y)\}.$$

Write $\Omega_n = \Omega^+_{01}(y_n)$. Notice that $A(\cdot)$, $B(\cdot)$ and $C(\cdot)$ are all right-continuous. We also know from Equations (5) and (6) that $A(\cdot)$, $B(\cdot)$ and $C(\cdot)$ are non-decreasing.

Since $y_n \downarrow y$, $C(\cdot)$ is a non-decreasing and right continuous, we have $C(y_n) \downarrow C(y)$, which implies that $\Omega_1 \subseteq \Omega_2 \subseteq \cdots \subseteq \Omega_n \cdots \subseteq \Omega$. Since $B(\cdot)$ is right continuous and non-decreasing, we know each Ω_n must take the form of $\Omega_n = [t_n, \overline{y}]$, where $\overline{y}$ is the upper boundary of $\mathcal{Y}$ (can be ∞). Let $\Omega^+_{01}(y) = [t, \overline{y}]$. Since t_n is a monotonically decreasing sequence and is bounded below by t, it must be a convergent sequence. Let t^* be its limit. If $B(\cdot)$ is continuous at t, it is straightforward to see that $t^* = t$. If t is a jumping point of $B(\cdot)$, since $B(\cdot)$ is right continuous, there must exists a N such that for all $n > N$, $t_n = t \Rightarrow t = t^*$. So regardless in which case we always have $t = t^*$. Therefore $t_n \downarrow t$.

Now we are ready to verify that $\inf_{y' \in [t_n, \overline{y}]} A(y') \to \inf_{y' \in [t, \overline{y}]} A(y')$. Since $A(\cdot)$ is non-decreasing, we must have $\inf_{y' \in [t_n, \overline{y}]} A(y') = A(t_n) \to A(t) = \inf_{y' \in [t, \overline{y}]} A(y')$, where the convergence holds by the right continuity of $A(\cdot)$.

The proof that $\tilde{g}_0^{-1}(y)$ is a well defined CDF can be similarly derived.

A.2. **Step 2: Show that $(\tilde{U}, \tilde{V})$ is a well defined copula.** Now we proceed and prove that $(\tilde{U}, \tilde{V})$ is a well defined copula. We begin by proving it is well defined subcopula.

Definition 1. *A two-dimentional subcopula (or brief subcopula) is a function* C *with the following properties (Nelsen, 2006):*

(1) *Domain(C)=$D_1 \times D_2$, where D_1 and D_2 are subsets of $[0, 1]$ containing 0 and 1.*
(2) $C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0$, *for all $u_1, u_2 \in D_1$ and $v_1, v_2 \in D_2$ such that $u_1 \geq u_2$ and $v_1 \geq v_2$.*
(3) $C(u, 1) = u$ *and* $C(1, v) = v$ *for all $u \in D_1$ and for all v in D_2.*

By Nelsen (2006), it is sufficient to show that $C(\tilde{g}_d^{-1}(y), p(z)) \equiv \mathbb{P}(\tilde{U} \leq \tilde{g}_d^{-1}(y), \tilde{V} < p(z))$ is a well-defined subcopula on $S_1 \times S_2$ where $S_1 = \cup_{(y_0, y_1) \in \mathcal{Y} \times \mathcal{Y}} \{\tilde{g}_0^{-1}(y_0), \tilde{g}_1^{-1}(y_1)\} \cup \{0, 1\}$ Property 1 and 3 holds straightforwardly by construction. It remains to verify property 2.

We consider the following cases:

(1) $\tilde{g}_1^{-1}(y_1') \geq \tilde{g}_1^{-1}(y_1)$ and $p(1) \geq p(0)$.

$$
\begin{aligned}
C(\tilde{g}_1^{-1}(y_1'), p(1)) - C(\tilde{g}_1^{-1}(y_1'), p(0)) - C(\tilde{g}_1^{-1}(y_1), p(1)) + C(\tilde{g}_1^{-1}(y_1), p(0)) \\
= \mathbb{P}(y_1 < Y \leq y_1', D = 1|Z = 1) - \mathbb{P}(y_1 < Y \leq y_1', D = 1|Z = 0) \geq 0.
\end{aligned}
$$

The last inequality holds due to equation (6).

(2) $\tilde{g}_0^{-1}(y_0') \geq \tilde{g}_0^{-1}(y_0)$ and $p(1) \geq p(0)$.

$$
\begin{aligned}
C(\tilde{g}_0^{-1}(y_0'), p(1)) - C(\tilde{g}_0^{-1}(y_0'), p(0)) - C(\tilde{g}_0^{-1}(y_0), p(1)) + C(\tilde{g}_0^{-1}(y_0), p(0)) \\
= \mathbb{P}(y_0 < Y \leq y_0', D = 0|Z = 0) - \mathbb{P}(y_0 < Y \leq y_0', D = 0|Z = 1) \geq 0.
\end{aligned}
$$

The last inequality holds due to equation (5).

(3) $\tilde{g}_1^{-1}(y_1) \geq \tilde{g}_0^{-1}(y_0)$ and $p(1) \geq p(0)$.

$$
\begin{aligned}
C(\tilde{g}_1^{-1}(y_1), p(1)) - C(\tilde{g}_1^{-1}(y_1), p(0)) - C(\tilde{g}_0^{-1}(y_0), p(1)) + C(\tilde{g}_0^{-1}(y_0), p(0)) \\
= [\mathbb{P}(Y \leq y_1, D = 1|Z = 1) - \mathbb{P}(Y \leq y_1, D = 1|Z = 0)] \\
- [\mathbb{P}(Y \leq y_0, D = 0|Z = 0) - \mathbb{P}(Y \leq y_0, D = 0|Z = 1)] = -H(y_0, y_1) \geq 0.
\end{aligned}
$$

The last inequality is proved in Claim 1.

(4) $\tilde{g}_0^{-1}(y_0) \geq \tilde{g}_1^{-1}(y_1)$ and $p(1) \geq p(0)$.

$$
\begin{aligned}
C(\tilde{g}_0^{-1}(y_0), p(1)) - C(\tilde{g}_0^{-1}(y_0), p(0)) - C(\tilde{g}_1^{-1}(y_1), p(1)) + C(\tilde{g}_1^{-1}(y_1), p(0)) \\
= [\mathbb{P}(Y \leq y_0, D = 0|Z = 0) - \mathbb{P}(Y \leq y_0, D = 0|Z = 1)] - \\
[\mathbb{P}(Y \leq y_1, D = 1|Z = 1) - \mathbb{P}(Y \leq y_1, D = 1|Z = 0)] = H(y_0, y) \geq 0.
\end{aligned}
$$

The last inequality is proved in Claim 2.

A.3. **Step 3: Show that $(\tilde{U}, \tilde{V}) \perp Z$.** Since by construction, for each tuple (y, d, z), the quantity $\mathbb{P}(\tilde{g}_d(\tilde{U}) \leq y, \tilde{V} < p(z)|Z)$ does not depend on Z. The independence assumption is satisfied straightforwardly.

A.4. **Proof of Claims.**

Claim 1. *If* $\tilde{g}_1^{-1}(y_1) \geq \tilde{g}_0^{-1}(y_0)$ *and* $p(1) \geq p(0)$, *then* $H(y_0, y_1) \leq 0$.

Proof. By definition of $\tilde{g}_d^{-1}$, we have

$$\begin{aligned}\tilde{g}_1^{-1}(y_0) - \tilde{g}_0^{-1}(y_1) = &\left[\mathbb{P}(Y \leq y_1, D = 1|Z = 1) - \sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0)\right] \\ &- \left[\mathbb{P}(Y \leq y_0, D = 0|Z = 0) - \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)\right] \\ &\equiv \mathcal{T}(y_1, \underline{y}_0(y_0), y_0, \overline{y}_1(y_1)) \geq 0.\end{aligned}$$

Case 1: $\sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0) \geq \mathbb{P}(Y \leq y_1, D = 1|Z = 0)$ and $\mathbb{P}(Y \leq y_0, D = 0|Z = 1) \geq \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)$. In this case, we have

$$\begin{aligned}H(y_0, y_1) = &\left[\mathbb{P}(Y \leq y_0, D = 0|Z = 0) - \mathbb{P}(Y \leq y_0, D = 0|Z = 1)\right] \\ &- \left[\mathbb{P}(Y \leq y_1, D = 1|Z = 1) - \mathbb{P}(Y \leq y_1, D = 1|Z = 0)\right] \\ \leq &\left[\mathbb{P}(Y \leq y_0, D = 0|Z = 0) - \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)\right] \\ &- \left[\mathbb{P}(Y \leq y_1, D = 1|Z = 1) - \sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0)\right] \\ &= -\mathcal{T}(y_1, \underline{y}_0(y_0), y_0, \overline{y}_1(y_1)) \leq 0.\end{aligned}$$

Case 2: $\sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0) < \mathbb{P}(Y \leq y_1, D = 1|Z = 0)$ or $\mathbb{P}(Y \leq y_0, D = 0|Z = 1) < \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)$. In this case, we have

$$\begin{aligned}&\mathbb{P}(Y \leq y_1, D = 1|Z = 0) > \sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0) \\ &\qquad \Rightarrow y_1 \notin \Delta_{01}^+(y_0) \Rightarrow H(y_0, y_1) < 0.\end{aligned}$$

or

$$\begin{aligned}&\mathbb{P}(Y \leq y_0, D = 0|Z = 1) < \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1) \\ &\qquad \Rightarrow y_0 \notin \Omega_{01}^+(y_1) \Rightarrow H(y_0, y_1) < 0. \quad \square\end{aligned}$$

Claim 2. *If* $\tilde{g}_1^{-1}(y_0) \geq \tilde{g}_1^{-1}(y_1)$ *and* $p(1) \geq p(0)$, *then* $H(y_0, y_1) \geq 0$.

Proof.

$$\begin{aligned}\tilde{g}_0^{-1}(y_0) - \tilde{g}_1^{-1}(y_1) &= \left[\mathbb{P}(Y \leq y_0, D = 0|Z = 0) - \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)\right] \\ &\quad - \left[\mathbb{P}(Y \leq y_1, D = 1|Z = 1) - \sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0)\right] \\ &\equiv \mathcal{H}(y_1, \underline{y}_0(y_0), y_0, \overline{y}_1(y_1)) \geq 0.\end{aligned}$$

Case 1: $\sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0) \leq \mathbb{P}(Y \leq y_1, D = 1|Z = 0)$ and $\mathbb{P}(Y \leq y_0, D = 0|Z = 1) \leq \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)$. In this case, we have

$$\begin{aligned}H(y_0, y_1) &= \Big[\mathbb{P}(Y \leq y_0, D = 0|Z = 0) - \mathbb{P}(Y \leq y_0, D = 0|Z = 1)\Big] \\ &\quad - \Big[\mathbb{P}(Y \leq y_1, D = 1|Z = 1) - \mathbb{P}(Y \leq y_1, D = 1|Z = 0)\Big] \\ &\geq \Big[\mathbb{P}(Y \leq y_0, D = 0|Z = 0) - \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)\Big] \\ &\quad - \Big[\mathbb{P}(Y \leq y_1, D = 1|Z = 1) - \sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0)\Big] \\ &= \mathcal{H}(y_1, \underline{y}_0(y_0), y_0, \overline{y}_1(y_1)) \geq 0.\end{aligned}$$

Case 2: $\sup_{y' \in \Delta_{01}^+(y_0)} \mathbb{P}(Y \leq y', D = 1|Z = 0) > \mathbb{P}(Y \leq y_1, D = 1|Z = 0)$ or $\mathbb{P}(Y \leq y_0, D = 0|Z = 1) > \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)$. Suppose the latter is true, that is,

$$\mathbb{P}(Y \leq y_0, D = 0|Z = 1) > \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1)$$

We will argue that it must be the case that $y_0 \in \Omega_{01}^+(y_1)$. We prove it by contradiction. Suppose $y_0 \notin \Omega_{01}^+(y_1)$. Since by the previous argument (as in step 1) we know $\Omega_{01}^+(y_1)$ take a form of $[t, \overline{y}]$, then $y_0 \notin \Omega_{01}^+(y_1) \Rightarrow y_0 < t$. Since $\mathbb{P}(Y \leq \cdot, D = 0|Z = 1)$ is right continuous and non-decreasing, it is necessary that $\inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1) = \mathbb{P}(Y \leq t, D = 0|Z = 1)$. Since $y_0 < t$, we must have

$$\mathbb{P}(Y \leq y_0, D = 0|Z = 1) \leq \inf_{y' \in \Omega_{01}^+(y_1)} \mathbb{P}(Y \leq y', D = 0|Z = 1),$$

which is a contradiction. Therefore we must have $y_0 \in \Omega_{01}^{+}(y_1)$, or alternatively by the definition of $\Omega_{01}^{+}(y_1)$, $H(y_0, y_1) \geq 0$.

By similar argument, we can show that if $\sup_{y' \in \Delta_{01}^{+}(y_0)} \mathbb{P}(Y \leq y', D = 1 | Z = 0) > \mathbb{P}(Y \leq y_1, D = 1 | Z = 0)$ is true, we must have $H(y_0, y_1) \geq 0$ too. □

Claim 3. *Inequality (3) holds.*

Proof. We restate Inequality (3) here:

$$\mathbb{P}(Y \leq y, D = 1 | Z = 1) + \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0 | Z = 1) \\ \leq \mathbb{P}(Y \leq y, D = 1 | Z = 0) + \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0 | Z = 0).$$

Note first that

$$\mathbb{P}(Y \leq y, D = 1 | Z = 1) - \mathbb{P}(Y \leq y, D = 1 | Z = 0) = \mathbb{P}(Y_1 \leq y, p(0) < V \leq p(1)) \tag{15}$$

On the other hand,

$$\begin{aligned} &\inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0 | Z = 0) - \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y \leq y', D = 0 | Z = 1) \\ &= \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y_0 \leq y', V > p(0)) - \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y_0 \leq y', V > p(1)) \\ &= \inf_{y' \in \Omega_{01}^{+}(y)} \left\{ \mathbb{P}(Y_0 \leq y', V > p(0)) - \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y_0 \leq y', V > p(1)) \right\} \\ &\geq \inf_{y' \in \Omega_{01}^{+}(y)} \left\{ \mathbb{P}(Y_0 \leq y', V > p(0)) - \mathbb{P}(Y_0 \leq y', V > p(1)) \right\} \\ &= \inf_{y' \in \Omega_{01}^{+}(y)} \mathbb{P}(Y_0 \leq y', p(0) < V \leq p(1)) \geq \mathbb{P}(Y_1 \leq y, p(0) < V \leq p(1)). \end{aligned}$$

where the last inequality holds by the definition of $\Omega_{01}^{+}(y)$ and Equation (2).

A similar result holds if we incorporate additional covariates X. □